

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 31 Mar 2003	3. REPORT TYPE AND DATES COVERED Final Technical Report, 15 Oct 02 – 31 Mar 03		
4. TITLE AND SUBTITLE Dynamic Measures of Spatial Ability, Executive Function, and Social Intelligence		5. FUNDING NUMBERS N00014-01-1-0091		
6. AUTHORS Dr. Gary Bradshaw, Principal Investigator Dr. J. Martin Giesen, Co-PI				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mississippi State University Department of Psychology P. O. Box 6161 Mississippi State, MS 39762		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Dr. Tanja Blackstone Office of Naval Research Code 34 800 North Quincy Street Arlington, VA 22217-5660		10. SPONSORING/MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Distribution Unlimited				
<div style="text-align: right; font-size: 2em; font-weight: bold;">20030613 137</div>				
13. ABSTRACT (Maximum 200 words)  By improving selection and classification of incoming military personnel, the Navy has an opportunity to recruit and retain talented sailors, thereby increasing readiness and reducing personnel costs. Here we introduce six new measures that are useful in personnel selection and classification. Three of the measures tap into spatial abilities useful in various contexts, such as radar operation and satellite imagery interpretation. The other three measures tap into executive functions, for example, the ability to focus one's attention on a visual task without being distracted by auditory information. This skill might be useful on the flight deck or other military situations where auditory distracters are loud and frequent. Our measures are based on objective behaviors, rather than subjective self-reports. As such, it is more difficult for subjects to intentionally bias their responses to put themselves in a positive "socially desirable" light. A series of six experiments were performed to establish the psychometric properties of the tests. Factor analyses of our experiments showed that the measures we developed are distinct from previously described ability tests. These measures also correlate with interesting demographic variables. Thus, they offer considerable promise for future personnel selection and classification				
14. SUBJECT TERMS Personnel selection; personnel classification; spatial ability;  executive function			15. NUMBER OF PAGES 5	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT  Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE  Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT  Unclassified	20. LIMITATION OF ABSTRACT  UL	

NSN 7540-01-280-5500

Computer Generated

STANDARD FORM 298 (Rev 2-89)  
Prescribed by ANSI Std Z39-18  
298-102

**DISTRIBUTION STATEMENT A**  
Approved for Public Release  
Distribution Unlimited

## FINAL TECHNICAL REPORT

GRANT #: N00014-00-1-0276

PRINCIPAL INVESTIGATOR: Dr. Gary L. Bradshaw (e-mail: [glb2@ra.msstate.edu](mailto:glb2@ra.msstate.edu))

INSTITUTION: Mississippi State University

GRANT TITLE: Dynamic Measures of Spatial Ability, Executive Function, and Social Intelligence

AWARD PERIOD: October 15, 2000 – September 30, 2002

OBJECTIVE: To develop new behavioral measures of spatial ability, executive function, and social intelligence that can be used for personnel selection and classification; to establish the reliability and validity of these new measures.

APPROACH: Six experiments were performed to explore new ability measures. All experiments were administered by computer and automatically scored. In the first experiment, 101 participants were administered a battery of 11 ability tests including 7 existing ability tests and 4 new tests we developed based on cognitive theories and research. The existing ability tests include the Revised Minnesota Paper Form Board Test, the ETS Identical Pictures Test, the Guilford-Zimmerman Spatial Visualization Test, Raven's Progressive Matrices Test, Thurstone's Perceptual Speed Test, the Integrating Details Test, and the Nelson-Denny Vocabulary Test. New tests include the Triesman Visual Search Test, the Change-Blindness test, the Shepard Mental Rotation Test, and the "Rule Discovery" Test. These tests were administered to paid subjects during three one-hour sessions. Factor analyses and clustering procedures were performed on various data sets to determine how the new ability tests relate to existing measures.

The Triesman visual search task requires participants to search a visual field for a target letter. In our version of this task, the target letter is always a T. Half of the visual fields contain distracter items consisting either of l's and Y's or l's and Z's. Previous research has found that participants are faster to locate the T with l and Y distracters than with l and Z distracters. In the former case, the T has a distinctive feature – the horizontal bar at the top of the letter, that sets it apart from the l and Y distracters, while no single feature distinguishes T from l's and Z's. Both accuracy and time measures are taken for each trial.

The Change-blindness test requires subjects to look at a picture that appears in two variations. One object is missing or moved between the two variations. The two variations alternate, but a neutral gray field appears between each picture, obscuring the change. After detecting the change, participants identify the locus of the change on the original image. Again, accuracy and reaction time data are collected for each trial.

The "Rule Discovery" measure is a new behavioral measure we have introduced. Participants taking this task are told they are supposed to discover a rule that maps strings, such as "~4J" into one of five categories. In fact, no rule governs the mapping between strings and categories. Instead participants receive a pre-determined sequence of "correct" and "incorrect" judgments. Regardless of their choice, every participant gets the first answer "correct." Regardless of their choice, the next six answers are "wrong." Performance rises from 30% in Block 1 to 80% in Block 4, a sign that success may be at hand. However, performance holds steady at 80% in Block 5, then drops abruptly in Block 6 to 40%, recovering to 70% by the final 8<sup>th</sup> block. At the end of each block we ask participants how well they did, how close they are to discovering the secret rule, and how well they anticipate performing in the next block. Although we have the classification of each item, the primary data in this task is how they answered the questions at the end of each block.

In the Mental Rotation task, participants view a figure made up from 9 blocks. Two views of the figure are presented at different planar rotations, and subjects must determine if the objects are identical or mirror-reversed. Accuracy and timing data are collected for various degrees of rotation.

The second experiment followed the same general pattern as the first: Five different measures were

administered to a group of 120 subjects. Four existing survey instruments were included where subjects report on their own characteristics. These include the NEOFFI personality inventory, the Rosenberg Self-Esteem Scale, the NPI Narcissism Scale, and the Rotter Locus of Control Scale. Participants also performed the rule discovery task. We also collected demographic information in the experiment, including participant gender, race, high school GPA, and ACT scores. Sessions lasted only an hour in length, and participants received class credit for their participation. An initial factor analysis was performed on the "rule discovery" data. Factors identified in this analysis were then correlated with the existing survey instruments and demographic information.

A third experiment replicated and extended the visual search measure introduced in the first experiment. Here we explored an unexpected finding in the visual search task, where participants were not as fast to identify distinctive items as in past research. Participants searched for the target letter T in a visual field that sometimes included 29 distracter items I and Y or distracter items I and Z. Sixty-four distinct stimuli were constructed. Subjects saw each stimulus display once per block for 8 blocks. The last block included 40 novel stimuli as well. Fifty-one participants received class credit for their participation in a single one-hour session. We compared the performance for "automatic" and "controlled" visual search problems across conditions using an analysis of variance.

Finally, three additional experiments were conducted on the Stroop effect and variations on this effect. The Stroop effect occurs when participants are given a color-naming task: to name the color of an object. Interference occurs when the object-to-be-named is a word that spells out a distinct color, such as the word RED printed in green ink. Interference also arises in a cross-modal version of the task, where participants hear the word "red" at the same time a patch of green appears on screen. A related phenomenon is known as negative priming. This effect appears across trials in a normal Stroop experiment as the color word on trial  $t$  becomes the color to be named on trial  $t+1$ . (An example, the word RED in green on trial  $t$  is followed by the word YELLOW in red on trial  $t+1$ .) In these cases, responses in the second trial will be slowed compared to responses to a different color. The first two experiments explored the use of the keyboard as a substitute for a voice key for negative priming and cross-modal Stroop effects, while the third experiment examined three versions of the Stroop task (normal, negative priming, and cross-modal). Demographic information was also collected in all three experiments, including information about typing skill. The first experiment included 47 participants who received class credit. The second experiment included 23 participants who received \$8 for participating in the one-hour session. The final experiment included 76 participants who received class credit for their one-hour session. The first two experiments were analyzed using t-tests, while a correlational analysis was performed on data from the final experiment to determine how cross-modal, negative priming, and ordinary Stroop effects related to one another.

**ACCOMPLISHMENTS:** Three new measures were tested in our first experiment: change blindness (the ability to identify a change in a visual scene); visual search; and mental rotation. The latter measure is sometimes used in a paper-and-pencil form on spatial ability tests, but our experiment indicates that the paper-and-pencil version only captures a part of mental rotation ability. In our first experiment, we demonstrated that:

- 1) Our new spatial ability measures were not strongly correlated with existing measures of spatial ability. This implies that we have not simply replicated a known measure in a new form, but instead are tapping into an ability that has not previously been recorded.
- 2) Our new spatial ability measures are not strongly correlated with one another, so they do not duplicate one another.

High performance in the change blindness test requires a strong ability to remember detail in a complex visual scene. This skill is needed in various investigative jobs, such as satellite imagery interpretation, and also in jobs such as radar operators, who need to track a dynamic set of existing objects and recognize new threats. We found large individual differences on this task: some individuals we tested could quickly recognize changes in the visual scene, while others never located the change. This ability, like others we identified, is not strongly correlated with indices of mental ability.

High performance in the visual search task requires people to pick out a non-distinctive target in a complicated visual background. Visual search is important in search-and-rescue missions, but could also be an important skill for urban combat, especially when combatants are disguised as civilians. Once again, we noted large individual differences in this ability, and it was not strongly correlated with intelligence.

Mental rotation requires the ability to manipulate three-dimensional objects mentally. This skill is useful for machinists, but is also likely to be important for pilots who maneuver in low altitudes. Once again, large individual differences were observed and were not correlated with intelligence.

A second experiment evaluated the potential of our "rule discovery" task. A factor analysis of the responses in the rule discovery task itself revealed four distinct factors. The first factor reflects how optimistic participants were in their ratings. Some participants predicted they would do well on the task, while others predicted poor performance. Factor 1 had a mild negative correlation ( $r = -.258$ ) with ACT score: participants who had high ACT scores may have realized the test could be difficult, and given a cautious prediction about their performance. Factor 2 is derived from the responses made when participants' performance is at its peak. Factor 2 is correlated with positive self-esteem on the Rosenberg scale ( $r = .263$ ) and also with an internal locus of control on the Rottenberg scale ( $r = -.241$ ). This factor appears to reflect a general optimistic tendency and a feeling that one is in control of one's fate. The third baseline performance factor is measured in the first two blocks of the session. This factor correlated significantly with the NEOFFI Openness scale ( $r = .335$ ), implying that participants who are "open to new experiences" rate themselves high at the beginning of the session, when compared to participants who are more closed in their attitude towards new events and opportunities. This factor also correlated with the math subscale of the ACT, indicating that people who have strong math abilities were likely to predict strong performance early in the test. The fourth and final factor is called the *setback performance* factor, as it depends on participants' reactions when their performance drops in the sixth and seventh blocks. This factor had a strong negative correlation with high school GPA: subjects who were upset by the setback tended to have higher high school GPA's, while subjects who were nonplussed by the drop in performance had lower high school GPA's. This factor did NOT correlate strongly with ACT scores, and thus is not a measure of intelligence. Rather, the factor seems to measure a personality variable, reaction to a setback, that influences academic performance. The setback factor also correlated with the NEOFFI Agreeableness scale ( $r = .233$ ).

Our third experiment explored an unexpected finding that appeared in the first experiment as we measured visual search times. Past researchers have shown that visual search can be quite fast if the target object differs from all distracter objects by a single distinctive feature, with little effect of the number of distracters on search time. Although we observed the general advantage of distinctive feature searches being faster than conjunctive feature searches, we still noted a strong effect of the number of distracters. Experiment 3 explored a possible explanation of this result: *differential memorization* of search items. The opportunity for differential memorization was present in our experiment as well as in previous research. However, we analyzed all stimuli, including the early trials when memorization could not have an effect, while other experiments had dropped early trials from the analysis, increasing the risk of contamination from differential memorization. Curiously, our third experiment failed once again to replicate classic research in this area. After 7 blocks of trials, we found no difference between the distinctive-feature items and the conjunctive-feature items. We now believe that the "automatic" fast search for distinctive items is only possible when the search is pre-programmed in advance, when participants are aware that a distinctive feature will be present and can plan for that event. Experiment three once again revealed strong individual differences in visual search ability.

Our fourth and fifth experiments were methodological in nature. Previous research had indicated that *keypads* were not effective in recording Stroop negative-priming effects and cross-modal Stroop effects. However, Stroop effects can be measured with a keyboard response (such as typing the first letter of the color name). If Stroop measures are to be practical, they must not rely on voice key responses, which means each participant must be run individually. Thus, our next two experiments explored whether keyboard input could be used to collect data about these effects. We were able to show that we could measure cross-modal and negative priming effects with a keyboard response, and that these effects were comparable in size to those measured with a voice key.

The final experiment exploited these methodological advances to determine how the Stroop effects correlated with one another. Negative priming effects were strongly correlated with ordinary Stroop effects ( $r = .64$ ), suggesting a common mechanism underlies both effects. However, cross-modal Stroop effects correlated only weakly with normal Stroop effects ( $r = .24$ ). Apparently some of our participants are able to ignore auditory information that is conflicting with a visual task, while others are distracted by the auditory inputs. The ability to focus on visual information while ignoring auditory inputs is likely to be important in areas of high noise, such as the flight deck of an aircraft carrier. Other Navy jobs might have the opposite needs. A sonar operator, for example, might need to be sensitive to auditory inputs even while working on apparently unrelated tasks.

In all of our Stroop experiments, typing skill was not related to the magnitude of the effect. It does not matter if our subjects are accomplished touch-typists or "hunt and peck" users. This will allow the test to be conducted even for recruits and sailors who have little or no typing skill.

**CONCLUSIONS:** Our six new measures are distinct from other ability measures (they do not replicate existing measures) and are different from one another (they do not duplicate each other). Several of the measures correlate with interesting personality characteristics, even though our tests are behavioral in nature. Further, all measures were designed for rapid administration and scoring. None of the measures takes more than 20 minutes to obtain from participants. Scoring is done as a part of the testing procedure, so that a report is available as soon as each participant completes each task. As such, these measures are promising new tools to employ for Navy recruit selection and classification.

**SIGNIFICANCE:** The new measures we have introduced have considerable potential as new ability measures useful in personnel selection and classification. We have already documented the novelty and practicality of the measures. It is also important to realize that our measures are all based on *behaviors* rather than on self-reports. Behavioral measures have an important advantage over survey instruments because they are less susceptible to bias or intentional manipulation on the part of test-takers. People are known to present themselves in a positive, socially desirable way on many survey instruments, especially when the instrument is being used to determine job assignments or promotions. The purpose of our behavioral measures is not at all obvious to most participants, and it is even more challenging to determine how behavior relates to the final score. In the rule-discovery task, for example, we learn which subjects are complacent with mediocre job performance, and which are driven to do better. This information is gleaned through an analysis of subject responses in six different questions. Contrast that with a self-report instrument such as:

How do you react when your job performance declines? (circle one)

- |                              |                                |
|------------------------------|--------------------------------|
| a) It doesn't bother me much | c) I try considerably harder   |
| b) I try a little harder     | d) I work very hard to improve |

Knowing the military's emphasis on performance and that one's job assignment and salary depended on the answer to this question, people with little or no motivation to improve are likely to overstate their willingness to work hard and improve.

A second important feature of the new measures we have developed is that they can be easily administered and scored by computer. Indeed, our measures do not even require a particular operating system, working well on both Macintosh and Windows computers.

Next, our measures have been shown to be distinct from our existing ability measures, tapping into abilities that have been neglected or overlooked in previous research. As such, they complement existing measures of ability. We believe that these measures tap into abilities that will be useful to identify for military selection and placement, although further research will be useful in validating these measures in an operational context.

**PATENT INFORMATION:** No patents have been filed or are anticipated from this research.

#### AWARD INFORMATION:

Outstanding Undergraduate Research Award: Shields, I., & Bradshaw, G.L. (2002). Keyboard Responses and the Negative Priming of the Stroop Effect. 1<sup>st</sup> Annual Mississippi State University Psychology Research Forum. Starkville, Mississippi, April 26, 2002.

Best Graduate/Post-Graduate Poster Presentation: Shields, I.S., & Bradshaw, G.L. (2002). Practical Measures of Negative Priming and the Cross-Modal Stroop Effect. 53<sup>rd</sup> Annual Mississippi Psychological Association Convention. Biloxi, Mississippi, September 6, 2002.

#### REFEREED PUBLICATIONS:

Bradshaw, G.L., Shields, I.S., & Giesen, J. M. (in preparation). Mechanisms of the Stroop effect: Evidence from Cross-Modal and Negative Priming Experiments. To be submitted to: Journal of Experimental Psychology: Learning, Memory, and Cognition.

Bradshaw, G.L., Giesen, J.M., Turner, T.R., & Shields, I.S. (in preparation). Preprogrammed Automaticity in Visual Search. To be submitted to: Memory and Cognition.

Bradshaw, G.L., Pate, M., Giesen, J.M., Bell, C., Shields, I.S., & Abernathy, L.T. (in preparation). Performance Monitoring as an Executive Skill. To be submitted to Journal of Educational Psychology.

Bradshaw, G.L., Giesen, J.M., Shields, I.S., Abernathy, L.T., & Kemp, L. (in preparation). New Dynamic Measures of Spatial Ability. To be submitted to Military Psychology.

#### BOOK CHAPTERS, SUBMISSIONS, ABSTRACTS, AND OTHER PUBLICATIONS:

Bradshaw, G.L., Giesen, M., Leathen, Shields, I., & Steinman, B. (2001). Dynamic measures of spatial ability, social intelligence, and executive function. MAS IV – Military Personnel Research, Memphis, Tennessee, June 4, 2001.

Shields, I., & Bradshaw, G.L. (2002). Keyboard Responses and the Negative Priming of the Stroop Effect. 1<sup>st</sup> Annual Mississippi State University Psychology Research Forum. Starkville, Mississippi, April 26, 2002.

Bradshaw, G.L., & Giesen, J. M. (2002). Computer-based behavioral measures of spatial ability, attention, and personality. MAS V -- Military Personnel Research. Memphis, Tennessee, June 5, 2002.

Pate, M., Bradshaw, G., Giesen, M., Abernathy, T., Mitchell, K., Bell, C., & Turner, R. (2003). Executive Monitoring of Performance in a "Rule Discovery" Task. 49<sup>th</sup> Annual Meeting of the Southeastern Psychological Association, New Orleans, Louisiana, March 26, 2003.

Turner, R., Bradshaw, G.L., Giesen, M., Abernathy, T., Mitchell, K., Bell, C., & Pate, M. (2003). Automaticity and Visual Search. 49<sup>th</sup> Annual Meeting of the Southeastern Psychological Association, New Orleans, Louisiana, March 26, 2003.

Shields, I., Bradshaw, G.L., & Giesen, M. Auditory and Visual Processes of Selective Attention. 49<sup>th</sup> Annual Meeting of the Southeastern Psychological Association, New Orleans, Louisiana, March 26, 2003.

Shields, I., & Bradshaw, G.L. (2003). Deconstructing Stroop: Individual Differences and Mechanisms of Various Stroop Effects. American Psychological Society, Atlanta, Georgia, May 31, 2003. (accepted)